DOCUMENT RESUME

ED 200 488 SO 013 251

AUTHOR Kemmis, Stephen
TITLE Program Evaluation in Distance Education: Against the Technologisation of Reason.
PUB DATE May 80
NOTE 32p.; A keynote address to the National Workshop on Distance Teaching sponsored by Townsville College of Advanced Education in co-operation with the Australian and South Pacific External Studies Association (Townsville, Victoria, Australia, May 11-14, 1980).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Correspondence Study; *Critical Thinking; Cultural Context; Educational Objectives; Elementary Secondary Education; Evaluation Criteria; Extension Education; Higher Education; Literature Reviews; *Models; Philosophy; *Program Evaluation; Trend Analysis
IDENTIFIERS Distance Teaching

ABSTRACT

Trends within the literature on program evaluation are reviewed and the implications of various trends for education are explored, with particular relevance for distance education. Distance education is an alternative educational approach based on correspondence, broadcasting, and long distance communication between teacher and student in a non-home based study study situation. Two major trends are identified--first, the trend toward the technologisation of reason (i.e., the acceptance of problem solutions by experts), and second, a trend toward concern with preserving a human dimension (i.e., convivality). With regard to educational evaluation, these trends have their counterparts in the move toward technical/quantitative assessment of performance and objectives and in the increasing interest in more individualized and responsive evaluation methodology. A new model is proposed to counteract shortcomings of these types of evaluation models and of other models in widespread use today (engineering models, ecological models, democratic models, illuminative/responsive models, etc.). This new model, based on evaluation as self-reflection and convivality, comes closer than any of the other models to patterning what people do in ordinary situations when they try to evaluate without the aid of specialists. Specifically, the evaluation as self-reflection model stresses rationality, autonomy and responsibility, the community of interests among participants in a specific program, value differences, explicit rules which govern procedures, and appropriate methods for evaluating the specific program or curriculum being studied. The conclusion is that this self-reflective evaluation model can help educators recognize one another's (and students) social being and develop a perspective which is more critical and reflective and less dominated by purely technical solutions. (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED200488

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Stephen Kemmis

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PROGRAM EVALUATION IN DISTANCE EDUCATION: AGAINST THE TECHNOLOGISATION OF REASON[1]

STEPHEN KEMMIS

DEAKIN UNIVERSITY

1981

SP013 251

PROGRAM EVALUATION IN DISTANCE EDUCATION: AGAINST THE TECHNOLOGISATION OF REASON[1]

Stephen Kemmis

Deakin University

In these remarks, I should like to address a very general problem of our culture as it comes to bear in the specific question before us today - the question of program evaluation in distance education. The general problem concerns the technologisation of our thought - our willingness to subjugate our human, convivial intelligence and capacity for critical thought to technologies of reason. In a world dominated by contending specialisms, arcane technologies and the proliferation of mass solutions to practical problems, we seem all too ready to give up our own critical intelligence and to accept as the solutions to our own problems the forms of life fashioned for us by others — where these "others" are "experts", "specialists" or, as I should prefer to call them for our purposes here, "the technologists of reason".

Let me set the stage by quoting the introduction to Stanley Aronowitz's (1977) paper "Mass Culture and the Eclipse of Reason: The Implications for Pedagogy":

> *In his book The Eclipse of Reason, Max Horkheimer, founder of the Frankfurt Institute for Social Research, provided one of the most succinct formulations of the problem engendered by mass culture. According to Horkheimer, the significance of the challenge posed by the massified culture industry to civilisation as such consisted in its assault on the capacity to engage in critical thought as a meaningful form of social discourse. Horkheimer cared deeply about the content of critical thought, but with the rise of fascism he became more concerned with the spectre of the end of reason itself. In his view, the capacity of humans to distance themselves from the object in order to gain critical perspective upon their social world can no longer be taken for granted. The restricted language and thought codes produced by the reduction of all thought to its technical dimensions reach far into the culture, encompassing schools as well as communications, the public as well as the private spheres of discourse. It is no longer a question of whether ordinary discourse is able to deal effectively with issues of specific ideological and social content. As Jurgen Habermas expressed it, the new situation raises the issue of the competence of people to effectively communicate ideational content. The issue is the capacity for theoretical or conceptual thought itself. When people lack such competence, social action that transcends the struggle for justice within the empirically given rules of social organisation and discourse is impossible. (p. 768)*

---

1 A keynote address to the National Workshop on Distance Teaching sponsored by Townsville College of Advanced Education in co-operation with the Australian and South Pacific External Studies Association, Townsville, May 11-14, 1980.

The problem to which Aronowitz addresses himself is that of our capacity to think critically about our social world. Technology has outrun the capacity of the ordinary person to understand; the legislative complexity and bureaucratisation of the administrative procedures of our society deny the layman access to understanding of access to broad arenas of social organisation. We have become, it seems, the instruments of the very social procedures and techniques we invented in order to instrumentatise the world to our needs and our desires.

It is no solution to hope that these cultural trends can be reversed: only the romantic believe that the way out of our cultural predicament is backwards. Fundamentalist religious zest and "small is beautiful" technologists cannot restore a presumed former state of innocence. There is no way back into the Garden. The technology awaits its use; the cultural desires which created it, once awakened, seek their fulfilment.

The scientists who developed nuclear technology during the Second World War faced personal crises because they knew that its capacity, once discovered, could not be covered up again (Jungk, 1960). Our choices seem not to be about the usefulness of the technology (for better or for worse, it was designed to be useful), they are about its wise use: we must control technology through wise and prudent use of its resources. To use it wisely, we must develop new social, political and procedural resources for critique, the capacity to "out-think" our technologies and to see the limits of their utility. The "spaceship earth" image, coupled with a vigorous ecological movement, has enabled us to see how our technological development efforts have turned on us and against each other. The "green revolution" liberates the capacity to feed the world's population but creates monocultures; our economic systems do not permit equitable distribution of these food resources without destroying the profitability which makes overproduction possible. In order to comprehend and cope with these contradictions, we must develop ways of sharing knowledge which liberate the power of critique and ways of making decisions which allow us prudently to control their consequences. We require as much boldness in critique and the sharing of knowledge as is presently invested in the production of technologies and their use for profitable purposes.

These are urgent imperatives. In education, as elsewhere in society, we need to re-establish the means by which we can recognise one another's social being and develop a critical perspective from which we can more reflectively control our development as a society and as a culture, as well as our development as a political economy.

The problem, as Aronowitz points out, is that we have been so dominated by the technical problems of our technological age that we tend to be satisfied by purely technical solutions. These tend to be "mass solutions": abstracted, generalised, universalised, pragmatic strategies for solving problems of a given type or class. They may or may not be solutions to our own problems. More importantly, because they are "off the shelf" solutions, they may distract our attention from the reality of our own situation, leading us to see this or that aspect with which a given technology can cope, and ignoring the wholeness, the manifold reality of our own situation.

There is a place in *War and Peace* where Tolstoy is discussing the question of free will and determinism. He makes a point there which bears closely on our discussion:

Only in our conceited age of the popularisation of knowledge, thanks to the most powerful weapon of ignorance — the diffusion of printed matter — the question of the freedom of the will has been put on a level, on which it can no longer be the same question.

In distance education, we are, of course very much involved in the business of the diffusion of printed matter (Tolstoy would undoubtedly have said the same for our audio and videotapes, and the impending computer revolution in distance education) and so there is a parallel between our primary task and the evaluation of that task. On the one hand, there is the trend to mass culture and the technologisation of reason; on the other, a concern for conviviality, for preserving a very human dimension in our affairs. In distance education there has long been a concern for humanising the processes of long-distance communication, for preserving the intimacy of the relationship between the learner and the teacher. At its worst, this has been regarded as a purely technical matter: a matter of motivation only, not of regard for the human dignity of both actors in the educational encounter. In contrast to this technical view, Ivan Illich (1971) proposed a system of "learning webs" (or networks, or exchanges) whose aim was to preserve conviviality while also maintaining the possibility of education at a distance. These were: reference services to educational objects, skill exchanges, peer-matching, and reference services to educators-at-large. There was considerable interest in those revolutionary ideas in the early '70s, but they seem to have succumbed to the power of another image — the mass culture image — of what education is and should be. They are now frankly dismissed by many as hopelessly romantic and in any case as more appropriate to Third World educational problems, not our own.

Perhaps the mere mention of that name — Ivan Illich — is irritating. Surely, some will say, those disquieting images and disturbing possibilities were laid to rest six or seven years ago. Or were they? It is my belief that the countervailing tendencies to technologisation on the one hand and conviviality on the other are as strong today as in the late 1960's and the early 1970's. As much as we may be concerned about the systems we have developed, we are concerned that we can recognise ourselves and one another in them as persons, not merely as instruments of economic processes.

This contrast between technologisation and conviviality has its counterpart in educational evaluation. On the one hand, there are trends abroad for efficiency in the management of mass education systems which represent the technologising view in educational evaluation. The development of a program for national assessment of educational performance in Australia is a stark manifestation of that trend. On the other hand, recent developments in evaluation methodology like Stake's (1975) "responsive" evaluation and Parlett and Hamilton's (1976) "illuminative" evaluation express an interest in conviviality, manifested in their use of the verstehen method which aims at empathetic understanding. They invite the more widespread use of interpretative methods (as opposed to the empirical-analytic methods of the "engineering" or "agriculture-botany" approaches).

The literature of "program evaluation" (though perhaps many do not now grace their present evaluation efforts with that graceless name) reveals that those two tendencies are chafing against one another within the field and in the work of program evaluation. Many may think of program evaluation as pure technology and as technologised. Many writers in the field are guilty, I believe, of promoting a technologisation of our reason in advocating models of evaluation based simply on inadequate models of

rational management. Liberal and gentle souls thrust into organisational evaluation settings, they choose the carapace of rational management as protection. Spare a thought for evaluators: their judgements make real people vulnerable and have real consequences for organisations. It is sometimes an uncomfortable role (though a good deal more comfortable than the role of the evaluator). Some evaluators find it safer to choose the dispassionate, objectivist methodology of the rational management approach to blunt the knife-edge of conscience. Others believe in the possibility of a purely rational society in which utilitarian ethics can ultimately advantage all.

Still other evaluators are willing to find more human solutions and to try the human way: they see conviviality as a goal, even in conflict, and propose less rationalistic methods. They demonstrate that conviviality in program evaluation is possible.

I hope to show that it is possible in program evaluation to resolve some of these issues. Technologisation of reason in the name of objectivity is not enough; conviviality by itself can be purely romantic notion. Beyond conviviality is the flintier, grimmer prospect of social criticism; "flintier" and "grimmer" not because the prospect of social criticism is grey, but because the task is a harder and more demanding one. It is the task of rebuilding institutions on the shifting sands of political alliances and defining their images in the mists of ideology, of contending rhetorics and aspirations. As such, it entails not only organisational change but cognitive change. As the Chilean physiologist Humberto Maturano (1974) has argued, any attempt to change society requires us to change the way we live; and this requires changes in both the cognitive and cultural domains.

Before proceeding further, it may be as well to make one important note with respect to program evaluation in distance education. While in many respects distance education differs markedly from conventional education, in questions of program evaluation, it is generally similar to conventional education. To be sure, there are differences in the character of the programs to be evaluated, especially in relation to the "delivery systems" and modes of communication with students, but there are nevertheless strong similarities: distance education course developers work together as conventional course developers do, the institutions offering distance education are educational institutions, and so on. In what follows, I have chosen to speak about program evaluation in general; I believe that my arguments apply as much to program evaluation in distance education as to more "conventional" forms of educational provision.

PROGRAM EVALUATION

It is by now more or less customary in evaluation to distinguish between <u>evaluation</u> and <u>assessment</u>. Assessment of student performance — evincing a concern for the outcomes of a course, curriculum or whole program -- is only a small part of evaluation. Beyond student assessment, I would like to direct your attention to at least the <u>evaluation of student learning</u> (the evaluation of the opportunities for learning provided by a course, curriculum or program), <u>curriculum evaluation</u> (the evaluation of the whole set of educational, social administrative and physical arrangements made by teachers and course developers for students), and <u>program evaluation</u> (which includes the former types of evaluation but expands them into the realm of the whole institution providing these educacional arrangements).

To put it very crudely indeed, evaluation might be considered at four levels, each including subsequent levels within its purview: (1) program evaluation, concerning general institutional arrangements; (2) curriculum evaluation, concerning the educational arrangements of whole curricula and particular courses; (3) the evaluation of student learning, concerning the opportunities for learning created by a particular teaching/learning encounter; and (4) student assessment, concerning the outcomes of student learning.

At the most general level, there is a range of alternative models of evaluation. The relationships between levels in these models are blurred and often obscure, however: it would appear that most authors in the field have been guided in their model-building activities by an over-riding image of the activity of evaluation. For example, Knox (1972) has in mind an image of the director of an adult basic education program who wants to collect fairly comprehensive data about how his program works, who it serves, how they experience it and where it can be improved. He provides a list of factors to consider under each of seven broad sub-headings for an evaluation plan:

I. Background regarding the anticipated contribution of evaluation to program improvement
   A. Primary purpose of evaluation for this course
   B. Symptoms that indicated that the existing procedures were inadequate
II. The approach to the evaluation plan
   A. Ideas about evaluation to keep in mind
   B. Types of evaluation data to be collected
   C. Functions of the division that may relate to course effectiveness
   D. Probable methods of data collection and analysis
   E. Using the results.

Stake (1967) seems to have in mind the inquiring teacher or outsider observing a teaching/learning encounter, or a sequence of such encounters held together in the framework of a course. His "countenance" of educational evaluation consists of a matrix of antecedents, transactions and outcomes by intents, observations, standards and judgements; the whole being considered against the background of the rationale for the program:

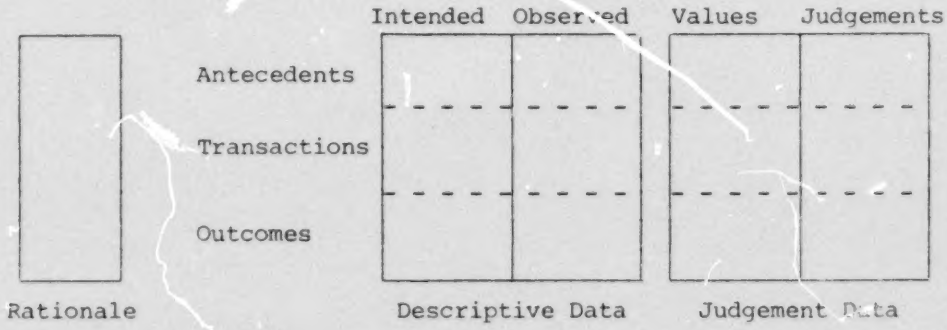

FIGURE 1: The matrix of Stake's "countenance" approach.

Stufflebeam, *et al.* (1971) have a model in mind of the rational manager of an input-throughput-output system which leads them to set out a range of considerations under each of the global headings of Context, Input, Process and Output (hence CIPP) which can guide the evaluator towards data likely to be relevant in making decisions about the whole program.

These examples serve only to illustrate the theme: that the evaluation of a whole program seems to be an extremely complex business, and that in order to reduce this complexity to manageable proportions specialists in the field have been guided by images of the task of the evaluator: harried program manager (in Knox's case), observer of educational encounters (in Stake's), or rational manager (or the servant of a rational manager, in Stufflebeam's).

For those interested in program evaluation in distance education, and especially interested in the evaluation of a distance education program as a delivery system, any of these three models might suffice. But, as my introductory remarks suggested, I think there are considerable dangers in treating an educational program as no more than a delivery system. By doing so, we may fall into the trap of technological reason: the trap of treating the delivery system as a technology which may be fine-tuned to improve its functioning or its efficiency, and thus of failing to subject it to substantive criticism as a form of life for students, teachers, ancillary staff and observers.

I would like to turn now to a consideration of five classes of evaluation approaches or, as I should prefer to call them, five images of evaluation which have some currency today. These are the engineering model, the organisational model, the ecological model, the illuminative/responsive model, and the democratic model. After considering each briefly, I would like to raise one or two general matters and then offer a sixth model, the model of evaluation as self-reflection, which I believe surmounts some of the difficulties of the other five. The sixth model sets out to incorporate the two notions discussed earlier: conviviality and critical reflection. As it happens, the model of evaluation as self-reflection is (I believe) also the model closest to what we do in ordinary situations when we try to evaluate an educational program without the aid of specialist approaches to the evaluation task.

THE ENGINEERING MODEL

The engineering model of educational evaluation is the one most familiar to curriculum and instructional developers. Writing about this model in 1979, David Jenkins, Barry MacDonald, Gajendra Verma and I had this to say:

> *Traditionally, educational evaluation has attempted to provide curriculum builders with some check on how well their intentions have been fulfilled in the educational performance of the programme under development. This has resulted (in general) in a concentration upon developers' instructional objectives and related student learning outcomes. This traditional model (which we call the 'engineering model' since it represents a technological solution to the problem of curriculum development) was first articulated by Tyler in 1934 and subsequently published as a curriculum for students of curriculum in 1949. In its most widely recognised form, it entails the following stages:*
>
> 1. *Secure agreement on the aims of the curriculum and/or course of study;*
> 2. *Express these aims as 'objectives' (i.e. explicitly stated student behaviours that the curriculum is intended to produce);*
> 3. *Devise and provide experience that seems likely to enable the learners to behave in the desired way;*
> 4. *Assess the congruence of student performance and objectives;*
> 5. *Vary the 'treatment' until behaviour matches objectives.*

> *This model, which integrates the evaluation and development processes, has exerted a powerful influence on curriculum evaluation on both sides of the Atlantic during the last twenty years. It has substantial merits of simplicity, rationality, utility and practicality. It specifies the task of the evaluator as that of measuring the extent to which the performance of students exposed to a particular educational treatment corresponds with the developer's expressed intentions. Almost all subsequent evaluation theorists have started from a consideration of this model, either extending and refining it, or reacting against it. (pp. 111-112)*

In a critique of the engineering model, we marshalled criticisms which were becoming widespread and grouped them under several categories:

(a) Problems of sufficiency

The engineering model is too narrow in its focus. Data about student outcomes are insufficient for making judgements about programs. Data about program circumstances and processes are as important as data about outcomes in reaching judgements of the overall worth of a program. Moreover, as Stake (1967) most cogently argued, a variety of values are relevant in reaching a judgement of its merit (not just those of the program developers), and judgements will vary with the values of the judge. His "countenance" paper argued that the plurality of perspectives and judgements needed to be taken into account in evaluation.

(b) Problems of specification

The behavioural specification of objectives required by the engineering model also proved unacceptable to evaluators who attempted to use it and go beyond it. Eisner (1967, 1969) proposed that "expressive objectives" should also be considered in curriculum, not just behavioural. Moreover, it became clear that teachers did not always follow developers' objectives, that they pursued many different objectives simultaneously in their teaching, and that different teachers pursue different objectives. Behavioural specification, while apparently acceptable in a closed world of the developers' objectives, did not meet the requirements of the real world of education. Scriven (1973) also mounted an attack on the objectives-based model, arguing instead for "goal-free" evaluation, in which the evaluator made a deliberate effort to avoid learning the program's goals and thus to escape the "tunnel vision" of looking only for the intended effects of the program (as distinct from unintended or side-effects of program operation). Finally, there is the problem that the specified objectives of the engineering model depend upon consensus about what is to be learned: in educational programs where broader social aims are pursued, such consensus cannot be assumed. Behavioural objectives may propose a consensus, but curriculum developers and evaluators cannot assume that specification will reduce division about social goals; on the contrary, it may increase it.

(c) Problems of measurement

The engineering model requires that criterion behaviours for a program can be specified; it also requires that the discrepancy between desired and actual achievement can be measured. But there are massive technical problems of developing tests of adequate validity and reliability for the program itself and the population to be tested. Walker and Schaffarzick (1974) demonstrated how often inappropriate tests have been used in

comparisions between innovative curricula and "conventional" alternatives, and how often the measured differences reflected the biases of the tests rather than the competences of students. Furthermore, in areas where broad attitudinal or social goals are being pursued by a program, tests are notoriously subject to bias — and the biases of the tests may or may not be the same as the biases of the programs under consideration.

(d) Problems of explanation

As Hastings (1966), Stake (1967) and Parlett and Hamilton (1976), among others have pointed out, the engineering model, with its emphasis on outcome measures, fails to give evaluators or curriculum developers an understanding of why a program fails when it fails. It fails to show how learners learn, or why they learn some things and not others. Most audiences of evaluation reports are interested in the conditions of learning, not just what is learned. Hence there has been a growth of studies of educational "transactions" (Rippey, 1973), the "social anthropology" of educational innovation (Parlett and Hamilton, 1976; Smith and Pohland, 1974), and the conditions under which learning takes place (Kemmis, 1977). I have addressed the problem more generally in proposing the wider use of "idiographic" methods in evaluation (Kemmis, 1978) in which close descriptions are made of the learning process in the evaluation of student learning.

(e) Problems of epistemology

The final, most general set of objections to the engineering model relates to its basis in a positivistic view of knowledge and learning. As we shall see later in this paper, this "objectivist" view of knowledge turns out to be narrowly scientistic (see Habermas, 1972), that is it is premissed on a view of science that regards the essential problem of the nature of truth as solved: according to this view, only technical problems in science remain; the nature of science itself is taken to be unproblematic. The danger is that the engineering model appears to treat evaluation data as value-free ("objective"). Yet, as Clarence Karier suggests, evaluation always has an ideological component; it always occurs within a value-framework:

> *It is my understanding that evaluation is a complex process of assigning values to phenomena, while ideology represents that set of values and attitudes which go to make up the composite picture of the social and individual philosophy by which men in a given culture profess to live. In this context, evaluation inevitably occurs within some kind of value orientation as part of an ideological framework.*

To some observers, the whole corpus of scientific knowledge as we know it is a uniquely, even peculiarly Western product, reflecting a faith that all problems — of production, of action, even of the soul — will yield to scientific investigation. Only recently have there been indications that this faith has been substantially dented (for example, in the failures of American military science and social engineering to win the war in Vietnam). But these dents show little sign of breaking that faith (as David Dawkins' paper at this seminar so ably demonstrates).

These problems with the engineering model might seem to have left it reeling, unable to attract champions in program evaluation today. But it is not so. The image which the engineering model embodies is a powerful one for our technological society, too powerful, perhaps, to be easily recovered from the store of our cultural archetypes and symbols, and subjected to critical scrutiny.

Eraut, MacKenzie and Papps (1975) have commented on the "mythology" surrounding educational development and its evaluation. Among the mythological aspects they identify are the following: that learning is to liberate the learner (on the contrary they point to the implicit authoritarian model of teaching and learning embodied in the standard view of educational development), the myth that students are the clients (in their view, teachers turn out to be the clients), the myth that the main *raison d' etre* of the paradigm is the improvement of students' learning efficiency (it turns out that students must adjust to the materials, not the other way around), the myth that student achievement is the all-important aim (it turns out that visible materials, not student achievement is what counts), the myth that knowledge is constructed like building with blocks (their observations led them to see the growth of knowledge as the student's reconstruction of a whole overlapping set of clusters of concepts and methods which constituted the discipline), and the myth that individualisation of learning improved learning achievement by giving the learner control over the learning process (their observation was that learners had control only over the pace and place of learning; in the tertiary course they studied, students demanded a more profound kind of control over their own learning — control over the cognitive organisation of the course).

These mythological elements of the technological approach to educational development identified by Eraut, MacKenzie and Papps suggest that the image of the engineering model has pressed deep into our understanding of educational development, but that the image is misleading. This is why it has a "mythological" quality. For this reason it may be asserted that the engineering model of educational development and evaluation represents a "technologisation of reason": it limits our thinking about these matters to technical and procedural dimensions. This reduction is fundamental to the model. It routinises the quintessentially human process of education so that it becomes more subject to technical management. In the next model to be discussed, the organisational model, this routinisation is projected onto an institutional plane: the interests of efficient management are extended to the management of institutions as social systems, subject to analogous technical thinking.

I have devoted so much space to the engineering model primarily because it is so profound a part of our culture, expressing so easily the technologisation of reason. It is, perhaps, the hardest image of program evaluation to extirpate from our thinking. Many evaluators have learned its limitations, however, simply by binding themselves to its dictates and then by considering how little of the life of the program they have captured in their methodological nets by applying it rigorously. There are others in the world of professional evaluation, though, who see things differently. The task to them is simpler; the world beyond the technology more dangerous. With a self-assurance rooted deep in our technological culture they model themselves on Ulysses, stopping the ears of their companions with wax and tying themselves to the mast of their methodologies. Thus they avoid, they think, the grasp of the sweet sirens of the wider world singing from the shore.

THE ORGANISATIONAL MODEL

The second image of program evaluation is based on the image of the program as an organisation. It is represented by the view of Knox (1972) discussed earlier (the harried program administrator attempting to gather data on the wide variety of concerns which might bear on a program). This model has become widely established in procedures for school accreditation, school self-study, and school review. The procedures have been further developed and refined in the context of evaluation in higher education.

Paul Dressel's (1976) *Handbook of Academic Evaluation* is representative of the genre.

The aim of the organisational model is essentially bureaucratic: to serve program managers (decision makers) with the range of information relevant in keeping their programs or institutions on the right track. The model is one of rational management, the image supported by cultural values of scientific rationality, economic efficiency, and consensus about an institutional mission: in short, the image is of the organisation man.

There is something peculiarly lulling about the organisational model: as each dimension of the institution covered by the model is mentioned, a litany of relevant information is recited. Rational management turns out to be a daunting task. So many things to be taken into account by the wise and prudent evaluator! There is not space here to recite the list in full, but it may be helpful to intone at least the macro-headings, the main dimensions, listed by Dressel in his consideration of institutional self-study (itself but one weapon in the arsenal of the institutional evaluator):

> A. *Determining institution purposes, goals and educational objectives (including a review of institutional documents, their consistency, the degree to which institutional subunits are in accord with these goals, resource provision, and the degree to which the goals are acknowledged by individuals and subunits throughout the institution)*
>
> B. *Measuring educational and other outcomes (including issues of what the outcomes are and whether all in the institution pursue them, the appropriateness of the outcomes to the clientele, the methods of assessment employed, institutional effectiveness in achieving other goals like community service and research)*
>
> C. *Evaluating learning experiences in terms of the desired outcomes Including issues of the types of experiences provided for students and the extent of their involvement in them, judgements of all relevant groups — teachers, students, administrators — and assessment of institutional climate)*
>
> D. *Evaluating the adequacy and utilisation of resources in terms of desired outcomes (including instructional staff, learning resources, financial resources and physical plant and equipment)*
>
> E. *Evaluating the planning and decision-making processes in terms of the desired outcome (including questions of who is involved in these processes, the groups beyond the institution who influence them, student involvement, and considerations of evidence related to gathering this data)*
>
> F. *Interpreting objectives, means of attainment, and evidence of attainment to new faculty, to students and to the public*
>
> *(pp. 419-422)*

There is a tremendous spuriousness and misplaced earnestness about such lists of relevant factors. One cannot deny their in-principle relevance to decisions about an institution, but the notion that any evaluation can achieve all of the data-gathering, analytic and reporting tasks set out is absurd. At best, the exercise is superficial and misleading, at worst, it is a mechanism for subordinating those involved to institutional goals which are expressed at a supra-individual, rhetorical level. The goal of "the examined life" may be worthwhile, but there must be time left to live it as well. Following the comprehensive prescription of an evaluation

system such as this is a sure way to get the inhabitants of the institution to leave evaluation to the experts. There is enough experience with such schemes in Australian schools for us to have some certainty that they tend to disaffect staff, to involve them in extensive data-gathering exercises which eat up scarce resources of time and money, and to yield little in terms of actual change.

It is worth noting, too, that the organisational model begins as the engineering model does: with a prescription of goals. All else hangs on this prescription. The same objections to goal-oriented evaluation apply here as to the engineering model.

Put simply, there are four kinds of problems with the organisational model: it takes a view of the institutional world as a complex set of interacting variables and is driven to absurd lengths to control this complexity; it is so comprehensive as to defy genuine application; it is politically insidious — it subjugates the critical perspective of inhabitants of the institution to the perspective of the institution as a corporate entity in its own right; and it is goal-based, and therefore open to some of the criticism made of the engineering model.

The organisational model I have presented here, the institutional self-study, has the redeeming features that it proposes the engagement of participants in a program in its evaluation, and that it suggests the enormous variety of places where something can go wrong with an organisation. The first is essential to good program evaluation, the second may be helpful. But as I have tried to show, too great an earnestness about following the prescriptions of the model will lead those involved away from the issues they can do something about. In this sense it exemplifies clearly the danger of the technologisation of reason: it may deaden critical awareness in the slow grind of its progress rather than heighten self-criticism in convivial debate.

THE ECOLOGICAL MODEL

A new model for educational evaluation is emerging out of the ecological psychological literature (e.g. Moos, 1976; Proshansky, 1970). A recent paper by Urie Bronfenbrenner, "The experimental ecology of education" (1976), points the way for a series of developments which may be quite interesting in program evaluation. Bronfenbrenner argues that programs are only inadequately tested in artificial or laboratory settings; that when and how people learn is affected by the relationships between learners and the surroundings in which they live and the relationships between the settings which constitute their environments; and that the methodology of choice must be what he calls "the ecological experiment". He employs the terminology (adapted from Brim, 1975) of micro-settings (the immediate setting containing the learner — e.g. home day-care centre), meso-systems (the interrelations between the major settings containing the learner at a particular point in his life — e.g. home, school, peer-group, television), exo-systems (an extension of the meso-system embracing concrete, formal and informal social structures — including the major institutions affecting the life of the child, like neighbourhood, mass media, agencies of government and informal social networks), and macro-systems (the overarching institutions of the culture or subculture — like economic, social, political and legal systems).

The image which underlies this model of program evaluation, if such it may be called, is the image of the evaluator as ecologist, the program as ecology. I mention it here because the image is a powerful one for us today. The model expresses a widespread desire in our culture to understand

our social life not just as a set of interacting variables, but as organic, structured and functioning. It sees a program as "alive", but more than that, as part of a "living" context.

Bronfenbrenner's paper makes interesting reading. The early results he cites in support of his approach are encouraging. It will be interesting to see whether these complex ideas are taken up widely and tested rigorously.

But these are not the methods for a program evaluator. They are methods for a social science researcher who has the leisure and the distance from a social reality to treat it as an object. In detail, the approach falls prey to the dangers of positivism cited in relation to the engineering model, and the problem of complexity raised with respect to the organisational model. (I do not mean that complexity by itself is a bad thing; rather, I mean to suggest that the kind of complexity these models deal in is hostile to understanding because it fragments our view of the program along the lines of its own analysis, not according to the problems perceived and experienced by those who inhabit programs as "worlds" in which they live).

These first three models (the engineering model, the organisational model and the ecological model) share a critical feature. All reveal an "objectivist" epistemology: the evaluator attempts to study the program "objectively" and dispassionately. All share an interest in value-neutrality with respect to the program. And all display what Habermas (1972) calls a "technical" knowledge-constitutive interest[2].

These approaches manifest most clearly the technologisation of reason, making critical self-understanding subordinate to program goals, bureaucratic organisational imperatives, or the "life" of the program as understood from a non-participatory (non-empathetic) perspective. The next three models to be discussed eschew the objectivist stance in favour of more empathetic and convivial approaches to understanding.

THE ILLUMINATIVE/RESPONSIVE MODEL

Recent evaluation literature has paid much attention to two relatively new models for evaluation: Parlett and Hamilton's (1976) "illuminative" approach and Stake's (1975) "responsive" approach.

Stake's responsive approach attempts to address issues raised by program participants rather than to import the questions of the evaluator into the program setting. He states:

> *An educational evaluation is responsive evaluation (1) if it orients more directly to program activities than to program intents, (2) if it responds to audience requirements for information, and (3) if the different value-perspectives present are referred to in reporting the success and failure of the program. In these three separate ways an evaluation plan can be responsive.*

2 Knowledge-constitutive interests are the human interests that motivate and guide the quest for knowledge. The empirical-analytic sciences (e.g. the physical sciences) are guided, Habermas argues, by "technical" interests, that is, interests in technical control of the phenomena studied; the hermeneutic or interpretative sciences (e.g. history) are based on a "practical" interest, that is, in understanding problems which have arisen and in accumulating experience which may guide us to act more wisely in future; and critical social science (e.g. political economy, some approaches in sociology) is guided by an "emancipatory" interest, an interest in emancipating people from the mystification of ideology, the dictates of

habit and precedent, or the compulsions of self-interested authorities. In deciding how to address the program, Stake's responsive evaluator is guided by the issues which concern those in and around the program. These issues guide the quest for data and the writing of reports. The evaluator is likely to choose to make portrayals of the program rather than report it in analytical terms. The aim is to convey something of the life of the program and the concerns of its inhabitants.

Similarly, Parlett and Hamilton's illuminative evaluation discards the analytical role in favour of an interpretative one. Parlet and Hamilton claim the authority of social anthropology for their view. On reflection, it seems that a wider interpretative tradition informs the methodology — the interpretative tradition of verstehen sociologists (see Outhwaite, 1975) and the hermeneutic or interpretative tradition in social science and history (see Kemmis, 1978). The illuminative evaluator chooses issues in a situation which help the reader to grasp its wholeness and to develop an empathetic understanding (verstehen) of the lives of those who inhabit it.

These methods, in my view, have much to recommend them. They set about deliberately trying to disclose the life of a social situation, progressively focussing on issues of particular concern to participants and accumulating data along the lines suggested by those issues.

The development of the illuminative and responsive models was a response to some of the perceived shortcomings of the engineering model. At a conference at Churchill College, Cambridge, in late 1972, U.S. and British "dissidents" from the engineering approach gathered to discuss the emerging alternatives. Among their number were Bob Stake (a psychometrician increasingly disenchanted by the empirical-analytic approaches to evaluation which characterised his own former work), Malcolm Parlett (an experimental psychologist who was similarly disenchanted by the limits of analysis and had turned to sociology and social anthropology for an alternative), David Hamilton (an ex-geology graduate, ex-teacher who had become interested in the social processes of classrooms), Mike Atkin (one of the prime movers in developing a high-school astronomy course in the U.S. and one of the earliest writers on the shortcomings of the objectives-based approach — see Atkin, (1963), Lou Smith (already attracting substantial attention for his ethnographic work in classrooms — see, for example, Smith and Geoffrey, 1968), Barry MacDonald (chief evaluator of the Humanities Curriculum Project in Britain, adopting a case study approach and concerned with the way evaluation could serve decision-makers and wider audiences of evaluation, and becoming sensitive to the politics of evaluation), David Jenkins (an ex-medieval religious drama scholar, ex-teacher, evaluator of the Keele Integrated Studies Project, a wit and raconteur whose brilliance as a writer emerges in his evaluation reporting), and a number of other names now well-known in evaluation circles.

Towards the end of the conference, a "manifesto" was drafted and endorsed by the participants. It was subsequently published in a "reader" on alternative evaluation methodologies, *Beyond the Numbers Game* (Hamilton, *et al.*, 1977):

> *On 20 December, 1972 at Churchill College, Cambridge the conference participants concluded a discussion of the aims and procedures of evaluating educational practices and agreed:*
>
> *I That past efforts to evaluate these practices have, on the whole, not adequately served the needs of those who require evidence of the effects of such practices, because of:*
> *(a) an under-attention to educational processes including those of the learning milieu;*

(b) an over-attention to psychometrically measurable changes in student behaviour (that to an extent represent the outcomes of the practice, but which are a misleading over-simplification of the complex changes that occur in students); and
(c) the existence of an educational research climate that rewards accuracy of measurement and generality of theory but overlooks both mismatch between school problems and research issues and tolerates ineffective communication between researchers and those outside the research community.

II *That future efforts to evaluate these practices be designed so as to be:*
(a) responsive to the needs and perspectives of differing audiences;
(b) illuminative of the complex organisational, teaching and learning processes at issue;
(c) relevant to public and professional decisions forthcoming; and
(d) reported in language which is accessible to their audiences.

III *More specifically they recommend that, increasingly,*
(a) observational data, carefully validated, be used (sometimes in substitute for data from questioning and testing);
(b) the evaluation be designed so as to be flexible enough to allow for response to unanticipated events (progressive focussing rather than pre-ordinate design); and that
(c) the value positions of the evaluator, whether highlighted or constrained by the design, be made evident to the sponsors and audiences of the evaluation.

IV *Though without consensus on the issues themselves, it was agreed that considered attention by those who design evaluation studies should be given to such issues as the following:*
(a) the sometimes conflicting roles of the same evaluator as expert, scientist, guide and teacher of decision-makers on the one hand, and as technical specialist, employee and servant of decision-makers on the other;
(b) the degree to which the evaluator, his sponsors and his subjects, should specify in advance the limits of inquiry, the circulation of findings, and such matters as may become controversial later;
(c) the advantages and disadvantages of intervening in educational practices for the purpose of gathering data or of controlling the variability of certain features in order to increase the generalisability of the findings;
(d) the complexity of educational decisions which, as a matter of rule, have political, social and economic implications; and the responsibility that the evaluator may or may not have for exploring these implications;
(e) the degree to which the evaluator should interpret his observations rather than leave them for different audiences to interpret.

It was acknowledged that different evaluation designs will serve different purposes and that even for a single educational programme many different designs could be used.

Looking back on the "manifesto" from the perspective of 1980, it is hard to remember how challenging those ideas were. They were prefigured by the work a considerable body of students of social and educational life, to be sure, but they represented an alliance of interests among a number of dissidents from the dominant evaluation paradigm on both sides of the Atlantic. The alliance continues to be a yeast to the dough of the evaluation literature.

The images represented by the illuminative and responsive approaches have venerable roots in social science, history and the humanities. Perhaps it is a British characteristic which simply lightens the consistency of the sometimes extremely turgid (and U.S.-dominated) evaluation literature, but the illuminative and responsive approaches seem more literate, wittier, and more artistic than much of the work reported in the "conventional" evaluation literature. Here we see the evaluator as social anthropologist, case-study worker, historian, portrayal-maker or artist. (Contrast the role of the evaluator as artist with the approach of Eisner who developed the notion of evaluation as art criticism; Eisner, 1974.)

These are images which not only attract but compel. They attempt to overcome the technologisation of reason with an art form. They attempt to heal the fragmentation of our understanding of a program produced by analytic (experimental, psychometric, objectives-based) methods and to provide holistic descriptions on the basis of which we, as readers, can judge the programs they portray for ourselves.

But the healing process is not so easily achieved.

The recipients of the first products of the "new wave" evaluation, as it came to be known (Stenhouse, 1975), must have been somewhat surprised by the sheer readability of the reports. But the response seemed short-lived. Very soon, the reports were being condemned as "subjective" and "impressionistic". It was argued that they could only be produced by a certain kind of person (the methods were so poorly understood that they could not be regarded as rigorous by audiences who had come to equate rigour with the rigidity of the prevailing technology). A long debate about the qualitative vs. the quantitative and the objective vs. subjective in evaluation began, though some might prefer not to describe it as a "debate"; opponents seemed to talk past one another as often as they bumped heads.

I have dwelt so long on the topic of illuminative and responsive models because they represent a substantial challenge to the engineering model in particular. And the challenge has been successful at least to the extent that the alternative approach has established itself as legitimate.

But several difficulties with illuminative and responsive approaches remain. The first is the problem that it is dismissed as "subjective" and "impressionistic". That these responses are based on ignorance of the methodological pedigree of the verstehen (empathetic understanding) and "hermeneutic" (interpretative; based on the scholarly tradition and history; see Gadamer, 1975) traditions need not concern us here. What should concern us is that this dismissal of illuminative/responsive evaluation reveals that the reports are not regarded as credible. This is, of course, a fundamental problem for an approach which is built on the notion that evaluation should speak to the concerns of participants and in their own preferred languages. I take it as an indicator of the extent of the technologisation of our reason that this failure to treat the reports of illuminative/responsive evaluation seriously should occur at all. Such is the imperialism of scientistic thought that it believes there are no truths to be told outside its own language. (How then do we justify history or criticism? As pastimes for effete intellectuals?)

The second difficulty is that posed by Walker (1977) in a paper entitled "Descriptive methodologies and utilitarian objectives: Is a happy marriage possible?" (The title almost describes the whole argument.) The problem is that descriptive methodologies, especially when they aspire to acute observation and profound insight, take time. But that is the commodity

program evaluators are usually short of. The "utilitarian objectives" of evaluation — getting the information to the decision-maker before the decision is to be made — are antipathetic to the descriptive objectives. Walker's argument may lend support to those who argue that the reports are "subjective" or "impressionistic" if it is read to suggest that they may be so because they are produced under pressure of time. (I am not personally convinced that this criticism is fatal to the project of illuminative/responsive evaluation; the process of "negotiation of accounts" to be discussed later in relation to the "democratic" approach, at least mitigates the force of the argument. Moreover, the process of writing up at least part of the life of a program, even if imperfectly, may help people to make wiser judgements of it; it is my view a crass mistake to believe that no knowledge is better than arguable knowledge about a program.)

A third difficulty is that these approaches tend to fall by one of the criteria they value most highly: the criterion of authenticity. The notions of "responsiveness" to participant and audience concerns, of "progressively focussing" on issues which compel those in and around the situation, and of "participant confirmation" (or "the shock of recognition") as guarantees of authenticity are challenged by the notion that it is not participants themselves but professional evaluators who shape the accounts. Perhaps this seems a purely technical problem: participants don't have the time to carry out the evaluation for themselves, don't have the appropriate skills, can't easily mediate with evaluation audiences with whom they have already-determined relationships. But it is not a merely technical consideration: the outside observer understanding and reporting on the program is doing something fundamentally different from the insider-participant who must live with the consequences of the evaluation, with the meanings reclaimed from its life, and the reputations the evaluation establishes or diminishes.

Though illuminative and responsive approaches are more convivial than their forbears in the evaluation literature of the last two decades, they have thisAchilles heel: that the life of the program is reconstructed for participants, not by them. In this sense, the illuminative/responsive evaluator still provides a technical service to the knowledge-generation processes of the program community. By specialising the knowledge-generation process with a band of professional evaluators, illuminative and responsive approaches preserve a crucial remnant of the technologisation of reason: someone else, not program participants or program audiences, knows best about how knowledge of the program is to be distilled and reported.[3]

THE DEMOCRATIC MODEL

The "democratic" approach to program evaluation attempts to provide a solution to this problem. It moves from the realm of processes of understanding social contexts to the question of the politics of information in evaluating them. In a sense, the shift from illuminative/responsive to democratic models parallels the shift from the engineering to the organisational model. Both the organisational and the democratic model incorporate

---

3 Note, however, that Stake and some co-workers have proposed that program participants develop their own evaluations or portrayals of their own situations. See, for example, Stake, *et al. The Accountability Notebook* (1971), Groteluesmen and Kemmis' (1973) notion of an evaluation file, and Stake and Scheyer's (1976) self-evaluation portfolio.

an elaborated view of the institutional uses of approach. In a widely-cited paper, "Evaluation and the control education", MacDonald (1976) distinguishes between bureaucratic, autocratic and democratic approaches to the evaluation of programs:

> Bureaucratic evaluation
> Bureaucratic evaluation is an unconditional service to those government agencies which have major control over the allocation of educational resources. The evaluator accepts the values of those who hold office, and offers information which will help them to accomplish their policy objectives. He acts as a management consultant, and his criterion of success is client satisfaction. His techniques of study must be credible to the policy-makers and not lay them open to public criticism. He has no independence, no control over the use that is made of his information and no court of appeal. The report is owned by the bureaucracy and lodged in its files. The key concepts of bureaucratic evaluation are 'service', 'utility' and 'efficiency'. Its key justificatory concept is 'the reality of power'.
>
> Autocratic evaluation
> Autocratic evaluation is a conditional service to those government agencies which have major control over the allocation of educational resources. It offers external validation of policy in exchange for compliance with its recommendations. Its values are derived from the evaluator's perception of the constitutional and moral obligations of the bureaucracy. He focuses upon issues of educational merit, and acts as expert advisor. His techniques of study must yield scientific proofs, because his power base is the academic research community. His contractual arrangements guarantee non-interference by the client, and he retains ownership of the study. His report is lodged in the files of the bureaucracy, but is also published in academic journals. If his recommendations are rejected, policy is not validated. His court of appeal is the research community, and higher levels in the bureaucracy. The key concepts of the autocratic evaluator are 'principle' and 'objectivity'. Its key justificatory concept is 'the responsibility of office'.
>
> Democratic evaluation
> Democratic evaluation is an information service to the community about the characteristics of an educational programme. It recognises value-pluralism and seeks to represent a range of interests in its issue-formulation. The basic value is an informed citizenry, and the evaluator acts as broker in exchanges of information between differing groups. His techniques of data-gathering and presentation must be accessible to non-specialist audiences. His main activity is the collection of definitions of, and reactions to, the programme. He offers confidentiality to informants and gives them control over his use of the information. The report is non-recommendatory, and the evaluator has no concept of information misuse. The evaluator engages in periodic negotiation of his relationships with sponsors and programme participants. The criterion of success is the range of audiences served. The report aspires to 'bestseller' status. The key concepts of democratic evaluation are 'confidentiality', 'negotiation' and 'accessibility'. The key justificatory concept is 'the right to know'.

The image of democratic evaluation thus poses a sharp challenge to the organisational model (which is essentially bureaucratic). The democratic evaluator stands outside the program as an independent, disinterested observer though

not an "objective" one: the "objective" outsider is serving "scientific" interests and thus adopts and interested, autocratic stance. He/she attempts to advocate the contending perspectives within it by proxy, fully and fairly, to regulate the processes by which participants come to agree that their perspectives are adequately represented (negotiation of accounts) and by which these views are gradually released to other participants in the situation (negotiation of release). The procedures for handling these negotiations are now susceptible of articulation and formulation as "principles of procedure" for evaluation studies (see Kemmis and Robottom, 1980).

The image of the democratic evaluator has its racier or more sinister side (depending upon your point of view): it is the image of the investigative journalist. This image is largely compatible with the image of illuminative or responsive evaluation, though perhaps it recognises itself as more "reactive": the investigative journalist shapes public affairs by reporting them; he/she does not merely report. (MacDonald, 1977, refers explicitly to the model of the "new journalist" as a source of techniques for the evaluator.) But it has another side too: that of the open-minded, concerned observer (based on the model of the "neutral chairman" of the Humanities Curriculum Project of which MacDonald was the evaluator; Jenkins, Kemmis and Atkin, 1977) who attempts to see that reason rather than political muscle wins the day as differing views of a program are shared among participants, and between participants and program audiences.

This latter part of the image of the democratic evaluation — as neutral chairman in a "discussion" about the program — is especially attractive. While in other respects the democratic approach is subject to the failings of the illuminative and responsive approaches, it is an attempt to overcome the political problem of authenticity: it attempts to respect the autonomy and responsibility and the reasonableness of program participants.

It is clear, however, that the democratic evaluator is still providing a technical service to a program community: that of "neutral chairmanship" in the debate about the program. Perhaps this is technologisation of reason: certainly the techniques for production and distribution of knowledge about the program are procedurally-bound. But I do not think that it can be so described: on the contrary, the democratic approach fosters, rather than denies, the capacity for critical thinking about the program. It thus comes closest of the images of evaluation so far discussed to achieving both conviviality and critical capacity.

Its failing, if such it may be called, is that the process of democratic evaluation as described by MacDonald is one of intervention into the social world of the program to provide this key service on behalf of program participants. Though it may provide a model of conviviality and criticism, it does not necessarily implant the model in the situation. As the notion is expressed ny Jurgen Habermas (1974), the ultimate problem is to create conditions within a program through which self-reflection may be established and sustained by participants on their own behalf. It is a problem of "the organisation of enlightenment", guided by an "emancipatory" interest (that of rising above the dictates of habit, precedent, coercion and material constraints towards greater reasonableness, self-awareness, discipline of thought and self-control and towards greater self-determination based on autonomy and responsibility and an awareness of the autonomy and responsibility of others). We will return to these notions in a later section.

THREE GENERAL ISSUES

Before proceeding, I should like to return to three general issues of concern in evaluation theory and method: the question of the objective and subjective

views of knowledge, the question of conviviality, and the question of autonomy and responsibility. By returning to these issues, I hope to emphasise several points made in my suvery of images of program evaluation so far.

(a) The question of objective and subjective views of knowledge

Many evaluators trained in social sciences have developed a perspective on the role and function of social science which depends on the notion of "objectivity": the capacity of science to free itself from commitments on one side or another of a value-question. Such evaluators hope that science can be value-free.

This position has come under heavy attack in the philosophy of social sciences in recent years (see, for just one example, Petrie, 1972). According to these writers, social science is not value-free: even the questions we ask are shaped by a framework of our own (or our culture's) values and interests. Habermas (1972) introduced the concept of "knowledge-constitutive" interests to describe the ways in which our inquiries are shaped to serve our values and our interests.

Maturana (1974) outlined the difference between the "objective" and "subjective" views in considering the question of the unity of mankind. According to the "objective" view, he suggests, men exist in an objective world that can be known and which exists independently of us as observers; our sense organs give us reliable knowledge of this world. The consequence of this view, he argues, is the view of mankind which sees us as genetically equivalent differing only in levels of endowment in physical or intellectual terms; which sees cultural differences as reflecting different modes of treating an objective reality; and according to which the cultural unity of man can only be obtained through the development of a culture in which men base their conduct on objective knowledge — that is, through the development of a culture in which men have "the right way" of looking at reality.

In contradistinction to the "objective" view, Maturana puts the view of cognition as a subject-dependent phenomenon. According to the "subjective" view, cognition as a process is bound constitutively to the organisation and structure of the knower; our history of interactions in the world defines the world for us, and it simultaneously defines us as "units of interactions" with certain potentials for interaction. The consequence of this view in relation to the cultural unity of mankind is that cultural differences do not reflect different modes of treating the same objective reality but legitimately different cognitive domains. Culturally different men live in different cognitive realities that are generated and specified by their living in them. The problem of the cultural unity of mankind, therefore, is not a problem of learning a single valid cognitive approach to objective reality (as rational empiricists would have it), but the problem of generating a common subject-dependent behavioural domain that defines a common subject-dependent reality. In short, Maturana argues that we should not expect to find the cultural unity of man by reference to a common objective reality "out there", as it were, but by living together so that we orient in common ways to the world and to one another.

This perspective is also implied by Wittgenstein's (1974) notions of "forms of life" and "language games". Truth, it may be argued, is not to be found entirely within language (empirical are about states-of-affairs in the world), nor is it to be found entirely outside language (it relates to propositions, not to experiences alone). Wittgenstein puts the problem thus:

> *So you are saying that human agreement defines what is true and what is false? — It is what human beings say that is true and false; and they agree in the language they use. That is not agreement in opinions but in form of life. (1974, p.88)*

The solution to the problem "how de we decide if a thing is true?" depends upon the way we use language, which in turn depends upon our having forms of life and sufficiently comprehensible to one another to support communication.

By discussing the question of the "objective" and the "subjective" in relation to the ideas of Maturana and Wittgenstein, I have hoped to show that the "subjective", the bane of "objective" social science, is not mere opinion, not mere individualism of perspective. Both the "objective" and the "subjective" depend upon agreements about how language is to be used. We have in fact arrived at an "interactive-subjective" perspective (not the lone subject experiencing and yet unable to test or to communicate that experience).

The images of program evaluation discussed so far may be arranged along this dimension from objective to subjective. The "objective" approaches underestimate the manifold quality of the world we experience, and they fragment our understanding of it. The most "subjective" image presented here (illuminative/responsive) is not yet fully conscious of its potential conviviality — its capacity to enter the program being studied, not to stand aloof from it as observer.

Many social scientists hope that the battle between the "objective" and the "subjective" can be won at that level. But some recent philosophers of social science do not think that this can happen. For example, von Wright in his book *Explanation and Understanding* (1971), argues that the two are simply incommensurable, that they are different modes which can never be reconciled; Habermas, in *Theory and Practice* (1974), however, argues that both are limited but they can be transcended in critical self-reflection. Choose your own unhappy ending.

This brief excursion into the world of the "objective" and "subjective" may convince you of little. But it may at least have raised the possibility that, to be convivial and to engage our understanding, program avaluation must have at least some element of the "subjective". Without the human process of expressing our experience in communication and orienting to one another in common action, we cannot hope fully to understand. This is the downfall of the "objective" evaluator, as it is of the evaluator who stands on the sidelines of the action as commentator or critic. As Mao Testung once put it: *"If you want to understand reality, you must participate in the struggle to change it".*

(b) Conviviality

Harry Wolcott[4], an ethnographer of education, was called upon to discuss a paper given at the Annual Meeting of the American Association for Research in

---

4 Wolcott's (1977) book *Teachers vs. Technocrats* bears greatly on our discussion — being both an ethnography (and thus a potential source of justification for advocates of the illuminative model of evaluation) and a treatment of the problem of lack of conviviality in program management and evaluation using the technology of PPBS (Program Planning and Budgeting System).

Education in San Francisco in 1976 by two evaluators of a technical/engineering/organisational cast of mind. These evaluators had defined education as "the collection of data to guide decisions to continue, revise or terminate programs". *"If you were an ethnographer"*, Wolcott remarked, *"how would you like your material to be used to continue, revise or terminate another culture?"*

The question is a pressing one for program evaluators: by adopting a technical role with respect to program participants the evaluator necessarily puts himself outside the life world of participants. Yet they must live with the consequences of the "truths" the evaluator tells.

In a recent paper on case study in research and evaluation (Kemmis, 1979) I argued that evaluators are, or seem to be, ruled by three "tyrannies":

> *(1) the tyranny that there is "a truth" to be told and that we can tell it,*
> *(2) the tyranny of conscience: that we believe we have a duty to arrive at judgements of the worth of social and educational programs, and to communicate these judgements — our judgements — to others, and*
> *(3) the tyranny of reporting: that we have a duty to commit these "truths" and judgements to paper and, through them, to create "the public meaning" (or at least "a public meaning") of the work.*

I see these tyrannies as evidence of the technologiation of our reason as evaluators, as evidence of what Sartre (1957) would describe as "bad faith" — our willingness to subjugate ourselves to a role thrust upon us sometimes by circumstance and sometimes by sponsors offering reputation or financial reward. (I am reminded of a Goon Show in which Eccles had been left to guard the door and keep out newspaper reporters while a secret meeting took place. The reporters managed to get into the meeting and Neddy Seagoon was demanding an explanation from Eccles. His reply? *"They burst through by forcing money into my hand".*)

The technologisation of our reason as evaluators and its attendant bad faith is made possible by the perspective we are invited to take on the program: we are observers, program participants the observed; we are the specialists, the program contains pathologies we are to diagnose; we are the therapists, participants must be demystified; we are the disinterested enquirers, participants are bound by self-deception and self-interest that only we can unravel and dispel. It is not surprising that we submit to the bad faith of the role: we are paid well for our work — in money, opportunities or reputation. And it is we, after all, who are the reputation-makers, with the power to distribute praise or blame, to confer the appellations "failure", "modest failure", "noble failure", "success".

Perhaps this characterisation seems extravagant or glib. But it is closer to the truth of the program evaluator's role than many would care or dare to admit. It is certainly the way in which we are perceived by a great number of those who suffer our services. And it is often the role we are enjoined to accept (quite deliberately) by evaluation sponsors — we can say harsh things and go away, the sponsors (who are very often the sponsors or administrators of the program too) must stay to live and rule another day.

Earlier, I pointed out that the value of authentic knowledge is that it provides a basis - the only certain basis - for program change. By "authentic knowledge", I mean knowledge grounded in the lives and circumstances of par-

ticipants. As Illich (1971), Freire (1970) and Habermas (1974) have been at pains to show, the capacity for critical self-reflection can only be developed in the language and perspectives of participants if the changes in a program are to "take" by conviction rather than coercion. As program evaluators outside other people's programs, we cannot tell only our truths; we must participate in the process of developing critical self-reflection alongside participants, as participants.

This is the way of conviviality. It is the development of critical self-reflection out of the talk and the (authentic) experience of co-participants. Language creates the possibility of critical thought (though it never guarantees it). The language of critical thought must recognise the experience and the language of those who want to participate in the critical debate; it must be self-reflection as well as critique of the social world of the program.

(c) Autonomy and responsibility

In any discussion of program evaluation, the question of autonomy and responsibility must arise. It is inextricably linked to the question of control of the program, since evaluation, whenever it aspires to influence program action, concerns program control.

Moral responsibility for an outcome can only be ascribed to a person to the degree that his or her free choice of action as an autonomous moral agent was a cause of that outcome. Just as the evaluator must assume that those involved in a program act rationally in the sense that they are open to arguments based on reason, so it must be assumed that they will act as autonomous and responsible moral agents.

This said, it becomes evident that a program is a community of self-interests (House and Care, 1977) based on the commitments of participants within the limits of their freedom, the opportunities and circumstances of the program setting, and the knowledge available about the consequences (intended and unintended, anticipated and unanticipated) of the program.

The autonomy and responsibility of those involved in a program is assumed by some of the images of program evaluation discussed here and negated by others. Indeed, it is a feature of the more "technological" models (especially the engineering model and the organisational model) that participants are treated as having subjugated themselves to the technology of the program. (In the use of the ecological model, it would appear that actors are assumed to be, to a greater or lesser degree of systemic constraints, determined rather than free to act as moral agents.) Program conditions which require this subjugation of participants may be coercive; program evaluators who assume this subjugation of participants are likewise coercive.

These three questions of "subjective" and "objective" views of knowledge, conviviality, and autonomy and responsibility must be actively considered in the development of a model of program evaluation capable of avoiding the limitations of the images of evaluation presented earlier. It is with these questions in mind that I now propose a sixth model of program evaluation: evaluation as self-reflection in the community of self-interests formed by a program.

EVALUATION AS SELF-REFLECTION IN A CRITICAL COMMUNITY

Jevons (1979) has remarked that the strategic aim of a tertiary education institution is that is become "a learning learning system", that is, a learning system capable of learning from its own experience. Schön (n.d.)

has also raised the general question "what is an organisation that it may learn?" He demonstrates that organisations do indeed learn from their experience.

Habermas (1974) has shown how this aim can be achieved in social organisations by the creation of conditions for "the organisation of enlightenment". He points out that three kinds of questions, each with its own criteria of adequacy, must be addressed in the organisation of enlightenment and the justification of action. First, at the level of scientific discourse, critical ideas which build towards a theory of the organisation must be developed. These must be tested against the usual canons of truth-telling. Second, at the level of the organisation of enlightenment itself, processes must be established which allow participants to engage in critical self-reflection on their own theoretical perspectives and practical commitments. These are to be judged by the criterion of free commitment to self-reflection and to the critical, co-operative project of establishing a consensus about social action based on mutual understanding and recognition of participants as persons in their striving for consensus. Third, at the level of practical action, processes must be established according to which practice can be improved and better adapted to the available opportunities and circumstances in the light of participants' intentions. Practical action must be judged by reference to the criteria of the wisdom and prudence of decisions about strategic action.

These three "levels" provide the substance for critical self-reflection. Kemmis and Hughes (1979) discussed evaluation as self-reflection in a critical community using these levels as substance for the discussions, and outlined an example of the approach in practice.

Beyond this substance, there is a need to create a climate in which self-reflective debate can be carried out. The appropriate climate is one described by the criterion for the organisation of enlightenment: that debate be carried out in an atmosphere of respect for persons and striving for consensus and mutual understanding. Kemmis (1980a) describes a strategy for developing the skills of "symmetrical communications" necessary to establish such a climate.

Seven principles for program evaluation describe the approach advocated here (Kemmis, 1980b). Together, they create an image of evaluation as the process of marshalling information and arguments which enables interested individuals and groups to participate in the critical debate (the process of self-reflection) about a program.

These principles are:

(1) *The principle of rationality as reasonableness*
Program participants act reasonably in the light of their circumstances and opportunities. It is the task of an evaluation to illuminate the reasoning which guides program development and evolution, to identify the contextual and historical factors which influence it, and to facilitate critical examination of these matters in and around the program community.

(2) *The principle of autonomy and responsibility*
Moral responsibility for an outcome can only be ascribed to a person to the degree that his or her free choice of action as an autonomous moral agent was a cause of that outcome. Curriculum development programs are co-operative enterprises in which individuals act as autonomous and responsible agents within limits of circumstance, context and (sometimes) coercion.

These factors limit individuals' accountability for program actions and their consequences. Evaluators must illuminate the interactive character of accountability for a program.

(3) The principle of community self-interests
When a curriculum development project is formed, it is a community of self-interests — it represents the self-interests of all participants within its terms of reference. The evaluator has a responsibility to illuminate the extent of commonality and conflict among the values and interests of participants in this community.

(4) The principle of plurality of value-perspectives
A range of different value-perspectives becomes relevant in judging a program. An evaluation should identify these different perspectives and be responsive to the different concerns they imply.

(5) The principle of the self-critical community: internal evaluation, evaluation consultancy, meta-evaluation, external and independent evaluation
Critical debate about the nature and worth of a program already exists within and around its program community. It is the task of program evaluation to refine this debate and improve its bearing on program action. Evaluation consultancy may provide additional tools for this purpose. Meta-evaluation efforts may help to improve the quality of the contribution of a program evaluation. An external evaluation may contribute to the critical debate by increasing awareness of a particular set of values and interests relevant to a program; it should not be thought of as an alternative to the self-critical process. An independent evaluation may help to harness program self-criticism where the program community is diffuse or divided by controversy. Self-criticism by the program community is the primary basis for program evaluation; other evaluation efforts extend it in different ways but do not supplant it.

(6) The principle of propriety in the production and distribution of information
Evaluation processes inevitably affect the political economy of information in a program (the production and distribution of information about it). Because information and arguments justify or legitimise decisions, evaluation affects the distribution of power and resources in program situations. Program participants and interested observers live with the consequences of use and abuse of evaluation information. An evaluation should have explicit principles of procedure which govern its conduct and its processes of information production and distribution.

(7) The principle of appropriateness
Evaluation design is a practical matter. An evaluation must be appropriate to the program setting, responsive to program issues, and relevant to the program community and interested observers. An evaluation design must be renegotiated as the study progresses in the light of changing circumstances, issues and interests, and in the light of its own consequences (as they become apparent).

This, then is the sixth and final image of program evaluation to be presented here. As distance educators, we must create the means to engage the community of participants in our programs (administrators, teachers, students and others) in the critical debate about them. More than this, we must refine the critical debate according to the principles outlined here. In short, we must make the community of the program a community of inquirers: a group jointly committed to developing the program as an expression of our common striving for mutual understanding and consensus about social action.

Some methods for pursuing this goal are available within the evaluation literature. Others are being developed in wider contexts: one is the notion of "action research" (see, for example, Corey, 1953) which I am currently investigating in a research project (funded by ERDC) with Robin McTaggart (Kemmis, 1980c). We are exploring the potential of these methods for creating critical communities of course developers, between teachers and students, and between schools and their communities. (Other collaborative research is being carried out in these areas at Deakin through the work of the Deakin University Action Research Group.)

The self-reflective approach is in some senses the most familiar notion of program evaluation. It is evident in a latent form whenever program participants discuss the life of a program with a view to understanding it in general, with a view to understanding it within the constraints of its context and circumstances (perhaps attempting to change its circumstances in order to overcome some of its limitations), or with a view to modifying program practice. In such circumstances, participants engage "naturally" in critical self-reflection on the principle of "concrete negation", that is, attempting to supersede the real constraints of present circumstances or present practice in the light of wider considerations.

Kemmis and Hughes (1979) proposed that this "naturally-occuring" form of self-reflection could be disciplined and extended by organisation of meetings of program participants for the primary purpose of critical self-reflection, as a conscious attempt to orchestrate the powers of critique in the program and thus to contribute to its theory, organisation and practice. Such meetings simultaneously contribute to the improvement of justification of the program and program action.

Self-reflective meetings provide a focus and forum for critical debate about the program (a basis for program critique). They do not supplant critical self-reflection elsewhere in the program; rather, they promote its general development. Moreover, self-reflective meetings can involve a range of different groups relevant to the program: teachers, students, administrators, sponsors, clients and interested observers can be invited to attend. In this way (cautiously and progressively) these groups can make their perspectives known and have them incorporated in (or excluded from) program justifications and action.

The meetings may have a slightly ritualistic quality in the sense that they do provide this forum and focus. A touch of formality may need to be present in the proceedings (a chairperson, explicit rules of order) in order that disparate views can be expressed and considered adequately. Nevertheless, the procedural aim must be to establish symmetrical communications as a way of working towards mutual understanding and consensus. Thus, part of the procedure must be to solicit perspectives across levels of the hierarchy of program organisation and administration (if it is hierarchically organised, as most programs are), and across degrees of participation (from "insider" to "outsider"). Naturally, the

meeting itself must have powers of decision which are binding on participants if it is not to be abused or co-opted by administrators as a means for pseudo-rational coercion or domination of participants (that is, subjugating the critical process to administrative self-interests).

The general problem with which this paper began was the cultural problem of the technologisation of reason. This technologisation undermines our capacity to engage in critical self-reflection. It is most evident in the engineering and organisational models for program evaluation, but the ecological, illuminative/responsive and democratic models are not yet complete in overthrowing its shackles. In one way or another, they limit authentic critique. At the general level, the technologisation of reason undermines our capacity to take a critical view of our own culture; at the particular level, it undermines our capacity to take a critical view of the educational programs we develop and of which we are members. The method of self-reflection in a critical community is an explicit attempt to create the conditions for authentic critique.

REFERENCES

ARONOWITZ, S. Mass culture and the eclipse of reason: The implications for pedagogy. *College English*, 1977, 8, 768-774.

ATKIN, J.M. Some education problems in a course content improvement project. *Journal of Research in Science Teaching*, 1963, 1, 129-132.

BRIM, O.G. Macrostructural influences on child development and the need for childhood social indicators. *American Journal of Orthopsychiatry*, 1975, 45, 516-524.

BRONFENBRENNER, U. The experimental ecology of education. *Educational Researcher*, 1976, 5, (9), 5-15.

COREY, S.M. *Action Research to Improve School Practices*. New York: Bureau of Publications, Teachers' College, Columbia University, 1953.

DAWKINS, D. Distance Teaching: corporate or collegiate? Paper presented at *ASPESA National Workshop on Distance Teaching*, Townsville College of Advanced Education, Townsville, Queensland, May, 1980.

DRESSEL, P. *Handbook of Academic Evaluation*. San Francisco: Jossey-Bass, 19 6.

EISNER, E.W. Educational objectives: Help or hindrance? *School Review*, 1967, 75, 250-266.

EISNER, E.W "Instructional and Expressive Objectives: Their Formulation and Use in Curriculum". *AERA Monograph Series on Evaluation*, No. 3 Chicago: Rand McNally, 1969.

EISNER, E.W. The perceptive eye: Towards the reformation of educational evaluation. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, 1974.

ERAUT, M., MacKENZIE, N. and PAPPS, I. The mythology of educational development: Reflections on a three-year study of Economics teaching. *British Journal of Educational Technology*, 1975, 6, (3), 20-34.

FREIRE, P. *Pedagogy of the Oppressed*. New York: Herder and Herder, 1970.

GADAMER, H.-G. *Truth and Method*. London: Sheed and Ward, 1975.

GROTELUESCHEN, A. and KEMMIS, S. *The Evaluation of Demonstration Library Projects*. Technical assistance provided for the Centre for the Study of Information and Education, Syracuse University, to the Division of Library Programs of the Bureau of Libraries and Learning Resources, U.S. Office of Education, 1973.

HABERMAS, J. *Knowledge and Human Interests*, trans. Jeremy J. Shapiro. London: Heinemann, 1972.

HABERMAS, J. *Theory and Practice*, trans. J. Viertel. London: Heinemann, 1974.

HAMILTON, D. *et al. Beyond the Numbers Game: A Reader in Educational Evaluation*. London: Macmillan Education, 1977.

HASTINGS, J.T. The why's of the outcomes. *Journal of Educational Measurement*, 1966, 3, 27-32.

HOUSE, E.R. and CARE, N. Fair evaluation agreement. Centre for Instructional Research and Curriculum Evaluation, University of Illinois at Urbana-Champaign (mimeo), 1977.

ILLICH, I. *Deschooling Society*. New York: Harper and Row, 1971.

JENKINS, D.J., KEMMIS, S. and ATKIN, R. "An Insider's Critique of UNCAL". In N. Norris (ed.) *SAFARI Papers 2: Theory into Practice*. Centre for Applied Research in Education, University of East Anglia, 1977.

JENKINS, D., KEMMIS, S., MacDONALD, B. and VERMA, G.K. "Racism and Educational Evaluation". In G.K. Verma and C. Bagley (eds.) *Race, Education and Identity*. London: Macmillan, 1979.

JEVONS, F. Possibilities for collaboration. *Open Campus*. (Occasional Papers published by the Centre for Educational Services, Deakin University), No. 1, Autumn, 1979.

JUNGK, R. *Brighter than a Thousand Suns*, trans. J. Cleugh. Hammondsworth, Middlesex: Penguin, 1960.

KARIER, C.J. "Ideology and Evaluation: In Quest of Meritocracy". In M.W. Apple *et al.* (eds.) *Educational Evaluation: Analysis and Responsibility*. Berkeley, Calif.: McCutchan, 1974.

KEMMIS, S. *How Do Students Learn? Working Papers on Computer Assisted Learning*. Occasional Publication No. 5, Centre for Applied Research in Education, University of East Anglia, 1977.

KEMMIS, S. Nomothetic and idiographic approaches to the evaluation of computer assisted learning. *Journal of Curriculum Studies*, 1978, 10, 45-59.

KEMMIS, S. Case study methods in research and evaluation. Paper presented at a conference in Case Study Methods in Research and Evaluation, Flinders University, Adelaide, March 6-7, 1979.

KEMMIS, S. Symmetrical communications: Developing mutual understanding and consensus in course teams. Paper presented at the *ASPESA Workshop on Evaluation in Distance Education*, Townsville College of Advanced Education, Townsville, May 11-14, 1980a.

KEMMIS, S. Seven principles for program evaluation in curriculum development and innovation. Deakin University (mimeo), 1980b.

KEMMIS, S. Research on action research: An interim report to the Education Research and Development Committee. Deakin University (mimeo), 1980c.

KEMMIS, S. and HUGHES, C. "Curriculum Evaluation in Higher Education: Self-Reflection in a Critical Community". In N. Unwin (ed.) *Research and Development in Higher Education 2*. T.E.R.C., University of New South Wales, 1979.

KEMMIS, S. and ROBOTTOM, I. Principles of procedure in curriculum evaluation. School of Education, Deakin University (mimeo), 1980.

KNOX, A.B. "Continuous Program Evaluation". In P.A. Taylor and D.M. Crowley (eds.) *Readings in Curriculum Evaluation*. Dubuque, Iowa: Wm. C. Brown, 1972.

MacDONALD, B. "Evaluation and the Control of Education". In D.A. Tawney (ed.) *Curriculum Evaluation Today: Trends and Implications.* London: Macmillan Education: Schools Council Research Studies, 1976.

MacDONALD, B. "The Portrayal of Persons as Evaluation Data". In N. Norris (ed.) *SAFARI Papers 2: Theory into Practice.* Centre for Applied Research in Education, University of East Anglia 1977.

MAO TSE-TUNG. *Selected Works of Mao Tse-Tung.* Peking: Foreign Language Press, 1974.

MATURANA, H. "Cognitive Strategies". In R. Abramovitz, *et al.* (eds.) *Cybernetics of Cybernetics,* B.C.L. Report No. 73.38, Biological Computing Laboratory, University of Illinois at Urbana-Champaign, May, 1974.

MOOS, R.H. *The Human Context: Environmental Determinants of Behavior.* New York: Wiley-Interscience, 1976.

OUTHWAITE. *Understanding Social Life: The Method Called Verstehen.* London: George Allen and Unwin, 1975.

PARLETT, M. and HAMILTON D. "Evaluation as Illumination: A New Approach to the Study of Innovatory Programs". In D.A. Tawney (ed.) *Curriculum Evaluation Today: Trends and Implications.* London: Macmillan Education: Schools Council Research Studies, 1976.

PROSHANSKY, H.M. *et al. Environmental Psychology: Man and His Physical Setting.* New York: Holt, Rinehart and Winston, 1970.

PETRIE, H.G. "Theories Are Tested by Observing the Facts: Or Are They?" In *Philosophical Redirection of Educational Research,* National Society for the Study of Education, 71st Yearbook, Part 1. Chicago: University of Chicago Press, 1972.

RIPPEY, R.M. (ed.) *Studies in Transactional Evaluation.* Berkeley, Calif.: McCutchan, 1973.

SARTRE, J.-P. *Being and Nothingness: An Essay on Phenomenological Ontology,* trans. H.E. Barnes. London: Methuen, 1957.

SCHON, D. What is an organisation that it may learn? Massachussetts Institute of Technology, (mimeo), n.d.

SCRIVEN, M. "Goal-Free Evaluation". In E.R. House (ed.) *School Evaluation: Politics and Process.* Berkeley, Calif.: McCutchan, 1973.

SMITH, L.M. and GEOFFREY, W. *The Complexities of an Urban Classroom.* New York: Holt, Rinehart and Winston, 1968.

SMITH, L.M. and POHLAND, P.A. "Education, Technology and the Rural Highlands". *AERA Monograph Series on Curriculum Evaluation,* No. 7. Chicago: Rand McNally, 1974.

STAKE, R.E. The countenance of educational evaluation. *Teachers' College Record,* 1967, 68, 523-40.

STAKE, R.E. *Evaluating the Arts in Education: A Responsive Approach.* Columbus, Ohio: Charles E. Merrill, 1975.

STAKE, R.E. *et al. The Accountability Notebook.* Centre for Instructional Research and Curriculum Evaluation, University of Illinois at Urbana-Champaign, 1971.

STAKE, R.E. and SCHEVER, P. A self-evaluation portfolio. Centre for Instructional Research and Curriculum Evaluation, University of Illinois at Urbana-Champaign (mimeo), 1976.

STENHOUSE, L. *Introduction to Curriculum Research and Development.* London: Heinemann Education, 1975.

STUFFLEBEAM, D.L. *et al. Educational Evaluation and Decision Making.* Itasca, Illinois: Peacock, 1971.

TOLSTOY, L. *War and Peace*, trans. Constance Garnett. London: Heinemann, originally published 1904; 1971 edition.

TYLER, R.W. *Constructing Achievement Tests.* Columbus, Ohio: Ohio State University Press, 1934.

TYLER, R.W. *Basic Principles of Curriculum and Instruction.* Chicago: University of Chicago Press, 1949.

von WRIGHT, G.B. *Explanation and Understanding.* London: Routledge and Kegan Paul, 1971.

WALKER, D. and SCHAFFARZICK, J. Comparing curricula. *Review of Educational Research*, 1974, 44, 83-111.

WALKER, R. "Descriptive Methodologies and Utilitarian Objectives: Is a Happy Marriage Possible?" In N. Norris (ed.) *SAFARI Papers 2: Theory into Practice.* Centre for Applied Research in Education, University of East Anglia, 1977.

WITTGENSTEIN, L. *Philosophical Investigations*, trans. G.E.M. Anscombe. Oxford: Basil Blackwell, 1974.

WOLCOTT, H. *Teachers vs. Technocrats: An Educational Innovation in Anthropological Perspective.* Centre for Educational Policy and Management, University of Oregon, Eugene, Oregon, 1977.